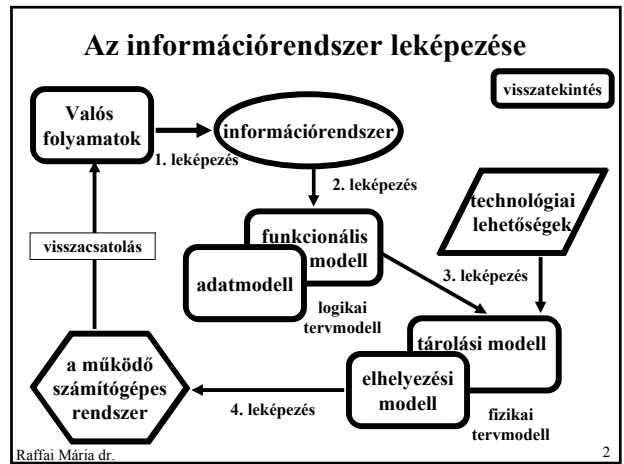




terminológia,
értelmezések,
fogalmak



Az adatmodell leképezése

Logikai komponensek	Fizikai komponensek
• egyedtípus	• állomány, file
• egyedelőfordulás	• rekord
• terjedelem	• rekordok száma
• tulajdonságtípus	• mező
• azonosító, egyedi ker. jell.	• elsődleges kulcs
• halmazkeresési jellemző	• másodlagos kulcs
• kapcsoló tulajdonságtíp.	• pointer, indextábla elem
• tulajdonság előfordulás	• mező értéke, adat
• terjedelem	• értelmezési tartomány
• adatmodell	• adatbázis

Raffai Mária dr. 3

- ### Az adategységek viszonya
- **Logikai egységek:**
bit → byte → mező → szegmens → rekord → blokk → file → adatbázis
 - **Fizikai egységek:**
 - belső ábrázolás szempontjából:
 - típus (integer, real stb.),
 - hossz (fix, változó)
 - I/O elérés szempontjából:
 - fizikai rekord
 - blokk
- Raffai Mária dr. 4

- ### Fizikai modellnézetek – adatmodell-leképezés
- **Elhelyezési modell:**
az a mód, ahogyan a háttértárolón elhelyezzük az adategységeket (fizikai rekord, blokkok, címkeinformációk, indexek, pointerok stb.)
 - **Tárolási modell**
azok a pontos információk, amelyek az egyes adategységek fizikai helyét határozzák meg (tárolótípus, tárolóközeg-azonosítás, fizikai címek stb.)
- Raffai Mária dr. 5

- ### Fizikai modellnézetek – a funkcionális modell leképezése
- **adatkezelési műveletek**
 - tárolóterület előkészítése
 - állományok létrehozása
 - állományok karbantartása
 - visszakeresési műveletek: teljes, blokk, bináris, egyedi keresés
 - **információelőállító műveletek**
- Raffai Mária dr. 6

Állománytípusok a funkcionalitás szerint (feldolgozás szerinti csoportosítási megközelítés)

- törzsállomány
- tranzakciós avagy átmeneti állományok
- lista- avagy jelentés-file-ok
- történet file-ok avagy archív állományok
- backup file-ok avagy biztonsági másolatok

Állományok az adatstruktúra belső megvalósítása szerint

- hierarchikus
- hálós
- asszociatív például B+ fa
- speciális architektúrák: hypertext, CAD/CAM file-ok



Hogyan helyezük el az
adatokat a
háttértárolókon?

(tárolási mód)

File-szerkezetek kialakítása (meghatározó tényezők, szempontok)

felhasználói igények
(feladatspecifikus
elvárások, platform-
független nézet)

tárolási és elérési lehetőségek
(platformfüggetlenség feltételek:
perifériák, adathordozók, input/
output vezérlő, adatbáziskezelők stb.)

file-szerkezet

Tárolás- és elérés-szervezés

- állományszervezés

→ adathordozóra-írás módja

- elérési mód

← visszaolvasás adathordozóról

- Fájlszervezési mód
az a rend, ahogyan az adatrekordokat s fizikai tárolóhelyeken elhelyezzük
- Fájl elérési mód (keresés)
az a lehetőség, ahogyan a tárolt adatrekordokat visszanyerhetjük
- Kapcsolatteremtés
az állományok rekordjainak egymáshoz rendelési módja

File-szervezési módok

- heap - szabad elhelyezés
- szekvenciális
 - fizikai és logikai szekvenciális
- indexelt avagy direkt
 - determinisztikus
 - egymáshoz rendelés: keresési kulcs - fizikai cím
 - algoritmussal: keresési kulcsból fizikai cím
 - indexelt szekvenciális: ISAM, VSAM, C_ISAM, B+ tree, R tree
 - random
- cluster szervezés (kapcsolat-file-ok)

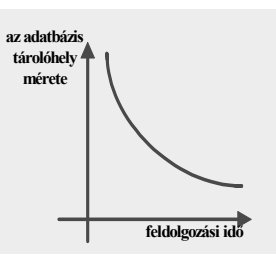
Keresési megoldások, visszaolvasás

Az a lehetőség, ahogyan a háttértárolón őrzött adatrekordokat feldolgozáshoz visszanyerhetjük.

Megoldások:

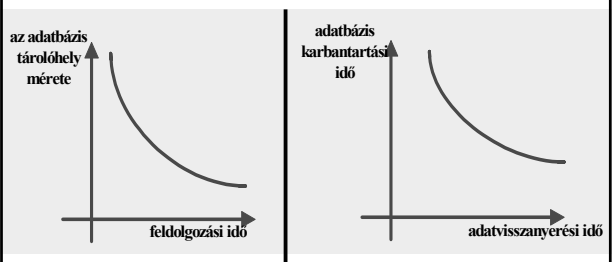
- fizikai cím ismerete nélküli keresés
- a rekord fizikai címe szerinti keresés
- keresés pointerláncokon keresztül
- keresés indextáblákkal

A tárolóterület és az időtényező viszonya



a feldolgozási idő a terület növekedésével csökken

A tárolóterület és az időtényező viszonya



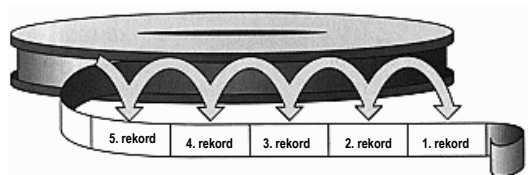
a feldolgozási idő a terület növekedésével csökken

hosszabb karbantartási idő - gyorsabb elérés

Néhány állományszervezési mód részletesebben (tárolási mód)

Szekvenciális file-szervezés

a rekordok fizikailag egymás után helyezkednek el, valamilyen szempont szerinti rendezett sorrendű feldolgozás biztosított)



felírási és visszaolvasási irány

Rekordok elérése szekvenciális állományokban

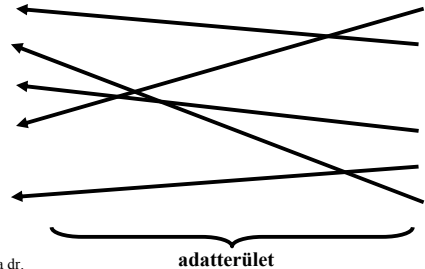
- **fizikai szekvenciális**
 - a fizikai sorrend azonos a logikaival
- **logikai szekvenciális:** a sorrendiség valamilyen technikával biztosított, csak közvetlen elérésű tárolón alkalmazható
 - teljes listaszervezet
 - indextáblában
 - másodlagos kulcs szerinti lánc
 - mutatótömbös megoldás

Raffai Mária dr.

19

Kezdőcím:
620

Logikai szekvenciális szervezés: teljes listaszervezet



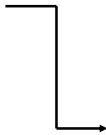
Raffai Mária dr.

20

Indextáblák alkalmazása rendezett beolvasáshoz

Adatállomány

Indextábla



Raffai Mária dr.

21

Indextömbök alkalmazása másodlagos kulcs szerinti eléréshez

Raffai Mária dr.

22

A szekvenciális file-szervezés

Előnyei

- gyors teljes keresés
- háttértároló-független
- jó tárolókihasználás

Hátrányok

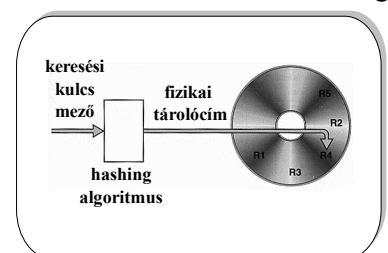
- egyedi keresés csak rekordvizsgálattal
- nehéz, lassú karbantartás
- körülményes kapcsolatmegvalósítás

Raffai Mária dr.

23

Direkt (indexelt) file-szervezés

kapcsolat van a rekord logikai azonosítója és az elhelyezés fizikai címe között
Csak közvetlen elérésű tárolón valósítható meg.



Raffai Mária dr.

24

Címzési technikák

- **Címzési technika**
 - abszolút cím
 - relatív cím: kezdőpont + eltolásérték
- **Direkt címzés**
logikai azonosító → fizikai cím közvetlen egymáshoz rendelése
- **Indirekt címzés** Fogalom: Túlcsordulás
logikai azonosító → többszörös leképezés (Hashing algoritmus) → fizikai cím

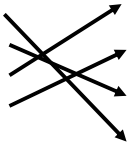
A direkt file-szervezés változatai

- **Közvetlen megfeleltetés**
- **Transzformációs eljárás** $\begin{cases} \text{determinisztikus} \\ \text{random} \end{cases}$
 - számítási eljárások: prímszámmal való osztás, alaptranszformáció, szorzás, négyzetre emelés, polinommal való osztás
 - helyérték kiválasztás
 - kombinált algoritmusok: hajtogatás, eltolás, csonkítás
 - nem numerikus azonosítók leképezése
- **Több szintű indextáblák alkalmazása**

Direkt file-szervezés indextáblával

Adatállomány

Indextábla



Leképezés valamilyen algoritmus szerint

Túlcsordulások kezelése

Két logikai azonosítóból ugyanaz a cím generálódik

Megoldás:

- újabb algoritmus alkalmazása
- független túlcsordulási területen elhelyezés
 - láncolási technika alkalmazásával:
 - egyedi rekordokat láncolunk vagy
 - bugyrot képezünk
 - láncolás nélkül bugyrok képzésével
- osztott túlcsordulási területen

A direkt file-szervezés

Előnyei

- gyors egyedi keresés
- gyors, egyszerű karbantartás
- egyszerű kapcsolat-megvalósítás

Hátrányok

- nehéz, lassú teljes keresés
- túlcsordulások problémája
- rossz tárkihasználás
- tárolóközeg-függő

Indexszekvenciális szervezés

egyesíti a szekvenciális és direkt szervezés előnyeit

Sajátosságai

- fizikailag folytonos, sorrend szerinti tárolás (azonosító, gyakoriság, betöltési sorrend)
- többszintű indextechnika
- különböző területek:
 - indexterület
 - elsődleges adatterület
 - túlesordulási terület
- törölt rekordok helyének felhasználása
- gyors egyedi és teljes keresés

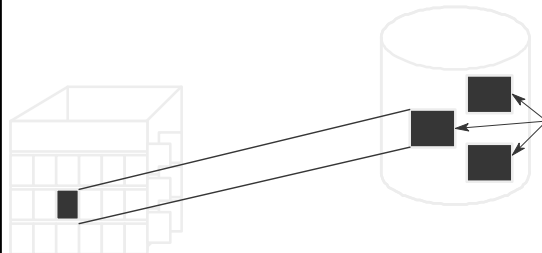
Indexszekvenciális szervezési módok

- ISAM: Indexed Sequential Access Method
- VSAM: Virtual Sequential Access Method
- C-ISAM
- B tree, B+ tree

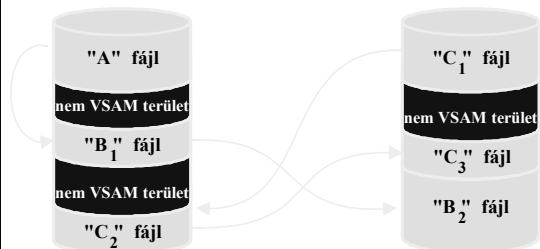
A VSAM sajátosságai

- virtuális tárolókezelés
- osztott terület alkalmazása
- ritka indexelési technika (nem minden rekordhoz tartozik indexbejegyzés)
- különböző szekvenciák kezelése
- dinamikus tárolókezelés
- változó hosszúságú rekordok kezelése
- sorrendhalmaz-indexek és osztott szabad terület alkalmazása

VSAM virtuális tárkezelés



A VSAM területmegosztása



Sorrendhalmazindexek
(indextábla sorai)

Vezérlő területek →

a sorok a vezérlő intervallumok

Szabad területek a vezérlő területen:

- osztott szabad terület (soronként)
- szabad vezérlő intervallum (önálló sorok)

A vezérlő intervallum feltöltése rekordokkal

1. Logikai rekord
2. Logikai rekord
3. Logikai rekord
-
- N. logikai rekord



szabad terület

- n. rekordleíró mező (RDF)
-
1. RDF
- vezérlő információk a terület- sávra



• Előnyök:

- fizikai adatfüggetlenséget biztosít **VSAM**
- a fizikai tárolás független a háttértároló típusától
- dinamikus helyviszanyerést valósít meg
- három különböző sorrendiséget tud kezelni → KSDS, ESDS, RRDS
- lehetővé teszi
 - a rekordok másodlagos kulcs szerinti keresését
 - katalógusrendszerével a központi nyilvántartást

• Hátrányok

- nem tud blokkolt rekordokat kezelni
- a programozó számára a rekordok láthatatlanok
- csak VSAM rutinokkal kezelhető → nem tesz eleget a nyitottság elvárásának

B tree

Jellemzői

- sűrű indexelési technika
- szekvenciális indextáblák (a logikai sorrend biztosítására)
- B+ fa indexek (a direkt keresésre), ezek többszintű indexállományok

Alkalmazás:

- relációs adatbáziskezelők
- programnyelvek pl. Pascal

A file-ok közötti kapcsolatok megteremtése

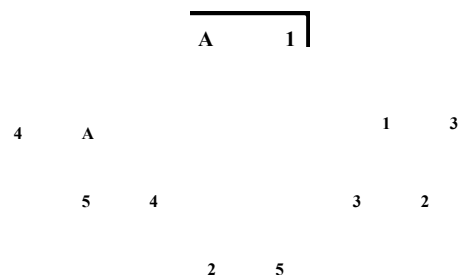
- rekordpárosítás azonos rendezettségű állományok között
- pointerek beépítése →
- pointerlánc: egyirányú, kétirányú →
- gyűrűs szerkezetek →
- pointer tömbök alkalmazása
- indextáblák létrehozása →

1:N fokú kapcsolat

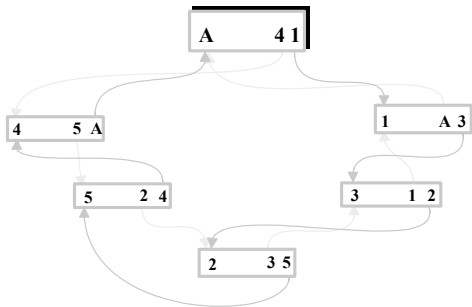
Láncszerkezet kialakítása pointerrel

Azonosító	Pointer	1	3
A	1	3	2
		2	E
Azonosító	Pointer	4	6
B	4	4	6
		6	E

Gyűrűs szerkezet egyirányú láncolással



Gyűrűs szerkezet kétirányú láncolással

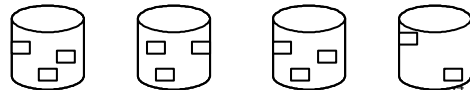


Állományon kívüli pointer hivatkozás

- Kapcsolat indextáblával



- Rekordfűzés cellás láncokkal



Állományszervezési mód megválasztása - működési hatékonyság

- tárolóterület nagysága
- logikai tárigény nem azonos a fizikaival
- időtényező: létrehozási, lekérdezési, karbantartási, feldolgozási idők
- hardver-szoftver támogatás: tárolók típusa, IOCS, adatkezelési lehetőségek

Állományszervezési mód megválasztása - felhasználói igények

- azonnali feldolgozások
- érvényességi kérdések
- eseményt követő feldolgozás, időbeliség
- feldolgozással szembeni elvárások
 - szekvenciális
 - közvetlen elérések gyakorisága
 - elsődleges, másodlagos kulcs szerinti keresések

Állományszervezési mód megválasztása - kiválasztási szempontok

- az adatbázis és állományainak jellemzői
 - adattömeg
 - azonosítók felépítése
- felhasználási paraméterek
 - tranzakcióarány - a feldolgozandó rekordok száma
 - feldolgozott rekordok aránya az összeshez képest
 - karbantartás gyakorisága, mértéke

Állományszervezési mód megválasztása - fizikai adatkezelési jellemzők

- központi tár mérete, címzési lehetőségek
- adatsatorna, puffereleési technika, kapacitás
- perifériavezérlő képessége, működési módja
- adathordozók jellemzői: címezhetőség, műveleti sebesség, kapacitás, cserélhetőség